```
              GenCore version 4.5
        Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        September 18, 2002, 22:26:45 ; Search time 8826.93 Seconds
                                              (without alignments)
                                              11851.432 Million cell updates/sec

Title:         US-09-645-593-8
Perfect score: 4999
Sequence:      1 ctccaagcatacggacaaggg........tcnattgtanatggncgttg 4999
Scoring table: OLIGO_NUC
               Gapop 60.0 , Gapext 60.0
Searched:      1799656 seqs, 1046326829J residues
Word size :    0
Total number of hits satisfying chosen parameters:      3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :     GenEmbl:*
               1:  gb_ba:*
               2:  gb_htg:*
               3:  gb_in:*
               4:  gb_om:*
               5:  gb_ov:*
               6:  gb_pat:*
               7:  gb_ph:*
               8:  gb_pl:*
               9:  gb_pr:*
              10:  gb_ro:*
              11:  gb_sts:*
              12:  gb_sy:*
              13:  gb_un:*
              14:  gb_vi:*
              15:  em_ba:*
              16:  em_fun:*
              17:  em_hum:*
              18:  em_in:*
              19:  em_mu:*
              20:  em_om:*
              21:  em_or:*
              22:  em_ov:*
              23:  em_pat:*
              24:  em_ph:*
              25:  em_pl:*
              26:  em_ro:*
              27:  em_sts:*
              28:  em_un:*
              29:  em_vi:*
              30:  em_htg_hum:*
              31:  em_htg_inv:*
              32:  em_htg_other:*
              33:  em_htgo_inv:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	C	Score	%	Length	DB	ID	Description
1		4999	100.0	4999	6	AX089358	Sequence
2		25	0.5	26541	2	AC095913	Rattus no
3		25	0.5	113725	2	AC108615	Rattus no
4	c	25	0.5	138733	9	AL136233	Human DNA
5		24	0.5	32	6	AX089372	Sequence
6		24	0.5	1683	8	AF240004	Sesamum i
7		24	0.5	133108	2	AC034196	Homo sapi
8		24	0.5	148598	2	AC022096	Homo sapi
9	c	24	0.5	152602	2	AC026193	Homo sapi
10		24	0.5	162433	2	AC024377	Homo sapi
11	c	24	0.5	163272	2	AC018496	Homo sapi
12		24	0.5	165077	2	AC026190	Homo sapi
13	c	24	0.5	169770	9	AC027763	Homo sapi
14	c	24	0.5	173878	9	AC040977	Homo sapi
15	c	24	0.5	176217	9	AC034192	Homo sapi
16		24	0.5	176219	9	AC087090	Homo sapi
17		24	0.5	191050	9	AC073565	Mus muscu
18		24	0.5	202704	10	AC090843	Mus muscu
19	c	23	0.5	29	6	AX089374	Sequence
20		23	0.5	420	5	AF013235	Quiscalus
21	c	23	0.5	606	5	AB067610	Ophthalmo
22		23	0.5	804	3	L09720 DDITC4RAN	Dictyosteli
23		23	0.5	1738	8	U43737 DMU43737	Drosophila
24		23	0.5	1905	3	X78119 PABTPRU1	P.amygdalus
25		23	0.5	4041	5	AF466307	Dictyoste
26		23	0.5	4644	9	AB017353	Rana rugo
27		23	0.5	42265	9	AC010504	Homo sapi
28		23	0.5	67687	2	AC101305	Mus muscu
29		23	0.5	72778	2	AC107045	Mus muscu
30		23	0.5	75572	2	AC101138	Mus muscu
31		23	0.5	77538	9	AC004500	Homo sapi
32		23	0.5	97057	9	AL627344	Human DNA
33	c	23	0.5	113725	2	AC108615	Rattus no
34		23	0.5	126910	9	AC026794	Homo sapi
35		23	0.5	128151	9	AP001966	Homo sapi
36		23	0.5	134757	9	AC007870	Genomic s
37		23	0.5	137336	8	AC007915	Genomic s
38		23	0.5	138752	2	AL161627	Human DNA
39		23	0.5	143799	9	AC011059	Homo sapi
40		23	0.5	147225	9	AC022498	Homo sapi
41		23	0.5	148898	9	AL121875	Human DNA
42		23	0.5	149133	9	HSBA51C14	Human DNA
43	c	23	0.5	150159	2	AC079089	Homo sapi
44		23	0.5	151301	2	AC018791	Homo sapi
45		23	0.5	154769	2	AL354717	Homo sapi
						AL671890	Mus muscu

ALIGNMENTS

```
RESULT (1)
AX089358
LOCUS       AX089358      4999 bp    DNA    linear   PAT 21-MAR-2001
DEFINITION  Sequence 8 from Patent WO0116340.
ACCESSION   AX089358
VERSION     AX089358.1   GI:13443626
KEYWORDS    .
SOURCE      flax.
ORGANISM    Linum usitatissimum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Malpighiales; Linaceae; Linum.
REFERENCE   1  (bases 1 to 4999)
AUTHORS     Chaudhary,S., van Rooijen,G., Moloney,M.M. and Singh,S.
TITLE       Flax seed specific promoters
JOURNAL     Patent: WO 0116340-A 8 08-MAR-2001;
            SemBioSys Genetics Inc. (CA) ; COMMONWEALTH SCIENTIFIC AND
            INDUSTRIAL RESEARCH ORGANISATION (AU)
FEATURES             Location/Qualifiers
     source          1..4999
                     /organism="Linum usitatissimum"
                     /db_xref="taxon:4006"
BASE COUNT     1340 a   1125 c   1192 g   1238 t    104 others
```

```
ORIGIN

Query Match      100.0%;   Score 4999;   DB 6;   Length 4999;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 4999;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0:

Qy    1 ctcaagcatacggacaaggtaaataacatagtcaccagaacataataacaaaagtgc   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 CTCAAGCATACGGACAAGGTAAATAACATAGTCACCAGAACATAATAACAAAAGTGC   60

Qy   61 agaagcaagataaaaaaaattagctatggacattcaggttcatattggaaacatcattatc  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGAAGCAAGATAAAAAAAATTAGCTATGGACATTCAGGTTCATATTGGAAACATCATTATC  120

Qy  121 ctagcttggacatccttcctcctgctctagttgagaggccttgggactaacgagaag     180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTAGTCTTGTGCGACCATCCTTCCTCCTGCTCTAGTTGAGAGGCCTTGGGACTAACGAGAAG  180

Qy  181 tcagttgggatagcagatcctatctgactagctagccttctggtgtttcagagtcttgt    240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 TCAGTTGGGATAGCAGATCCTTATCCTGACTAGCCTTTCTGGTGTTTCAGAGTCTTCGT   240

Qy  241 gccgccgtctacatctatctccattaggtctgaagatgactcttcacaccaacgacgttt   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GCCGCCGTCTACATCCATTAGTTCTGAAGATGACTCTTCACACCAACGACGTTT          300

Qy  301 aaggtcctatcctacctcagctgcaatacctgcctgcttgcagacctgcgagcatcgtgc   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 AAGGTTCTATCCTACCTCGAGCTGCAATACCTGCTTGCAATACCTGGAGCATCGTGC       360

Qy  361 acgatgattggatactgtgaggagagtgttgctgcgatttagagctcccggttgggtga   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ACGATGATTGGATACTGTGGAGGAGAGTGTTGCTGCGATTTAGAGCTCCCGGTTGGGTGA   420

Qy  421 tttgacttcgatttcagttaggcttgttgaaatttttcaggttccattgtgaagcctt    480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TTTGACTTCGATTTCAGTTAGGCTTGTTGAAATTTTTCAGGTTCCATTGTGAAGCCTT     480

Qy  481 agagttgagcttcctccatgtattgaaatcgaagtgcacattttttcaacgtgtccatcaatcca  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AGAGGTTGAGCTTCCTCCATGTTAATGCCTGAATCTCCTAGACAAAAGGGAAG          540

Qy  541 tcgatctgagtattgaaatcgaagtgcacattttttcaacgtgtccatcaatcca        600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TCGATTCTCGAGTATTGAAATCGAAGTGCACATTTTTTCAACGTGTCCAATCAATCCA     600

Qy  601 caaacaaagcagaagacaggtaatcttcatcttacttatacgacaagtaagtcttaccg    660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CAAACAAAGCAGAAGACAGGTAATCTTCATGACTATCCATAATCCATCAATCAATCCA     660

Qy  661 tcatgcataataacgtctcgttcctcaagaggggttttccgacatccataacgaccga    720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TCATGCATAATAACGTCTCGTTCCTTCAAGAGGGGTTTTCCGACATCCATAACGACCCGA   720

Qy  721 agcctcatgaaagcattaggaagaacttttggttcttctgtcatggcctttataggtg    780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 AGCCTCATGAAAGCATTAGGGAGAACTTTTCATTACTTATACTGACAAGTAATAGTCTTACCG  780

Qy  781 tcagccgagctgccaattccgtccgactggctccgcaaatattcgaacggcaagtta     840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 TCAGCCGAGCTGCCAATTCCCGTCCGACTGGCTCCGCAAATATTCGAACGGCAAGTTA    840

Qy  841 tggacttgcaaccataactcacggtattgagcaggacctattgtgaagactcatctcat   900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 TGGACTTGCAACCATAACTCCACGGTATTGAGCAGGACCTATTGTGAAGACTCATCTCAT   900

Qy  901 gggcttcagaatgtggttgtcagcaaaccaatgaccgaaatcatcacatgaccgacgt   960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 GGAGCTTCAGAATGTCGGTTGTCAGCAAACCAATGACCCGAAATGATCCCACCATGGC    960

Qy  961 ccagtgggtgagcgaaacagcaggaagcgcctatctttcagagtcgtgagctccacac  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CCAGTGGGTGAGCGGAAACGAACAGGAAGCGCCTATCTTCAGAGTCGTGAGCTCCACAC  1020


Qy 1021 cggattccggcaactacgtgttgggcaggcttcgccgtattagagatatgttgaggcaag 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CGGATTCCGGCAACTACGTGTTGGGCAGGCAGCCTTCGCCGTATTAGAGATATGTTGAGGCAAG 1080

Qy 1081 accatctgtgccactcgtacaattacgagagttgttttttgtratttcctaagttt      1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 ACCCATCTGTGCCACTCGTACAATTACGAGAGTTGTTTTTTGTGATTTTCCTAAGTTT    1140

Qy 1141 ctcgttgatggtgagctcatattctacacgtatggtctctcaacgtcgttcctgtcat 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CTCGTTGATGGTGAGCTCATATTCTACACGTATGGTCTCTCAACGTCGTTTCCTGTCAT  1200

Qy 1201 ctgatatccgtcatttgcatccacgtgcgcgcgcctccgtgcaagtcccctaggtgtca 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTGATATCCCGTCATTTGCATCCACGTGCGCGCGCCTCCGTGCCGTGCCCAGGTGTCA   1260

Qy 1261 tgcacgccaaattggtggtgtgcgggcgccctgccctgcttcttaccgatgggtgaggtt 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 TGCACGCCAAATTGGTGGTGTGCGGGCGCCCTGCCCTGCTTCTTACCGATGGGTGGAGGTT 1320

Qy 1321 gagtttggggtccgcggcgatgagtagtgggttgcggtttggtgtgggttgggttgacggca 1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 GAGTTTGGGGTCCGCGGCGATGAGTAGTGGGTTGACGGTTTGGTGTGGGTTGACGGCA    1380

Qy 1381 ttgatcaatttacttcttgcttcaaattcttggcagaaaacaattcattagattagaac 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 TTGATCAATTTACTTCTTGCTTCAAATTCTTGGCAGAAAACAATTCATTAGATTAGAAC  1440

Qy 1441 tggaaaccagagtgatgagacggattaagtcagatccaacagagttacatctcttaaga 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 TGGAAACCAGAGTGATGAGACGGATTAAGTCAGATTCAGATTCCAACAGAGTTACATCTCTTAAGA 1500

Qy 1501 aataatgtaaccccttagactttatatattgcaattaaaaaataatttaaacttttag 1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 AATAATGTAACCCCTTAGACTTTATATATTTATATATTTGCAATTAAAAAATAATTTAAACTTTTAG 1560

Qy 1561 actttatatatagtttaataactaagttaaccactctattatttatatcgaaactatt 1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 ACTTTATATATAGTTTAATAGTTTAATAACTAAGTTAACCACTCTATTATTATATCGAAACTATT 1620

Qy 1621 tgtatgtctccctaaataaacttgtattggtttacagaaccctataatcaaaatat    1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 TGTATGTCTCCCCTAAATAAACTTGGTATTGGTTTACAGAACCTATAATCAAAATAAT   1680

Qy 1681 caatactcaactgaagttgtcagttaattgaagggattaacggcaaaatgcactagt  1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 CAATACTCAACTGAAGTTGTGCAGTTAATTGAAGGGATTAACGGCCAAAATGCACTAGT  1740

Qy 1741 attatcaaccgaatagatcacactagatggccaatttcacatcaatatcatcgccgttctt 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 ATTATCAACCGAATAGATTCACACTAGATGGCCAATTTCCATCAATATCATCGGCCGTCTT 1800

Qy 1801 cttctgtccacatatcccctctgaaactgagagacaccgcacttcattgtccttatta 1860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 CTTCTGTCCACATATCCCCTCTGAAACTTGAGAGACTTGAGACGACAACGCACTTCATTGTCCTTATTA 1860

Qy 1861 cgtgttacaaatgaaacccatgcatccatgccaagaactgacggccaagaaccctttc 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 CGTGTTACAAATGAAACCCATGCATCCATGCCAAGAATGGCGCCAAGAACCCTTC      1920

Qy 1921 ccctccattcttattgtggcgccatccattcaccatctccgctaaaacacccccca   1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 CCCTCCATTCTTATGTGGCGACCATCCATTTCACCATCTCCCGCTATAAAACACCCCCA  1980

Qy 1981 tcacttcacctagaacatcatcactacttgcttatccatccaaaagataccccaccatggc 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 TCACTTCACCTAGAACATCATCATCACTACTTGCTTATCCATCCAAAAGATACCCACCATGGC 2040

Qy 2041 tagatcatcaagcccttgcttctctcactctgcatttcgccattctcctgcattcttc 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 TAGATCATCAAGCCCTTTGCTTCTCTCACTCTGCATTTCGCCATTCTCCTCCTTCCACTCTTC 2100
```

```
Qy  2101  tctgggtaggcagcaattccagcaggggaacgagtgccagatcgacagagatcgacgcatc  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  TCTGGGTAGGCAGCAATTCCAGCAGGGGAACGAGTGCCAGATCGACAGAGATCGACGCATC  2160

Qy  2161  cgagccggacaaacatccaggcagagagctggcaccatcgaggtatgggaccagaaccg    2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  CGAGCCGGACAAAACCATCCAGGCAGAGAGCTGGCACCATCGAGGTATGGGACCAGAACCG  2220

Qy  2221  ccagcaattccagtgcgcgtgtgtccgcttgtaaggcgcaccattgagcccaaaggtct   2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CCAGCAATTCCAGTGCGCTGCGTGTTGCCGTTGTAAGGCGCACCACCATTGAGCCCAAAGGTCT  2280

Qy  2281  tctcttgcctttctacagcaacaccctcagctcatctacatcgttcaagtataaatta    2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TCTCTTGCCTTTCTACAGCAACACCCTCAGCTCATCTACATCGTTCAAGTATAAAATTA  2340

Qy  2341  aatcagttcatacaatgataaccaccacttcgaatgtatttatcaaatatcaatgatcga  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  AATCAGTTCATACAATGATAAACCACCACTTCGAATGTATTTATCAAAATATCAATGATCGA  2400

Qy  2401  tgcaactgtatgtgttgtatattcagtaggggagttacagagggaatcatgttcccakga  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  TGCACCTGTATGTGTGTGTATTGTGTATTCAGGTAGGGGAGTTACAGGAATCATGTTCCCAKGA  2460

Qy  2461  tgtccagagacattgagggaatcccagcacgcaaggagcaacacaggggtagttcc      2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  TGTCCAGAGACATTGAGGAGAATCCCAGCAGCACAAGGACAACAGGGCCAACAGGGTAGTTCC  2520

Qy  2521  caagaccagcaccagaagatccgcgcttccgtgaaggtgacgtcattgccgtcctgcc    2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  CAAGACCAGCACCAGAAGATCCGCCGCTTCCGTGAAGGTGACGTCACTGCCGTCCCTGCC  2580

Qy  2581  ggtgtagcccactggtcctacaacgatggcaacgacccagtcatggtcattgtgtccat   2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  GGTGTAGCCCACTGGTCCTACAACGATGGCAACGATCCAGTCATGGTCATTGTGTCCAT  2640

Qy  2641  gacacttccagccacctcaaccaactggacaacaaccaacccccaggggtatataagcacgt  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  GACACTTCCAGCCACCTCAACCAACCTCAACAACAACAACAACCCCAGGGGTATATAAAGCACGT  2700

Qy  2701  agttgctaataaattgcacacaattggaactctattttcgtatctaataactttttcct  2760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  AGTTGCTAATAAATTGCACACAATTGCACACAATTGGAACTCTATTTTCAGTATCTAATAACTTTTTCCT  2760

Qy  2761  ttttggcagaacttctacttggcaggaaaccgagagacgagtcgaacaatgcagca     2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTTTGGCAGAACTTCTACTTGGCAGGAAAACCCGAGAGACGAGTCGAACAATGCAGCA  2820

Qy  2821  aggagcaggctgagccgtgggagagtgaagtgacgaggacgcaggggaacctcttca    2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  AGGAGCAGGCTGAGCCGTGGGGAGAGTGAAGTGACGAGGACGCAGGGGAACCTCTTCA  2880

Qy  2881  acctgcaacaacctcttcttgcggaatcgacttccggaggcgttcaatgt           2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  ACCTGCAACAACCTCTTCTTGCGGAATCGACTTCCGGAGGCGTTCAATGT  2940

Qy  2941  cgacgagaacgtggcaaggaggctacagagcgagaacgacaacagagccgagatcgtccg  3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  CGACGAGAACGTGGCAAGGAGGCTACAGAGCGAGAACGACAACAGAGCCGAGATCGTCCG  3000

Qy  3001  agtcgaagggcgagctcgacactgcgacctccgaccagtatccaggagagtcacagga    3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  AGTCGAAGGGCGAGCTCGACACTGCGACCTCCGACCAGTATCCAGGAGAGTCACAGGA  3060

Qy  3061  gcaggaggaggtcgtggtggtgccgcgctactactcctccatggaggtgagggagacctt  3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  GCAGGAGGAGGTCGTGGTGGTGGTGCCGCCGCTACTACTCCTCCAATGGAGGTGAGGGAGACCTT  3120

Qy  3121  catgagactaattgagaacatcggcgatccttctcggacagcatttttcactccagaagc  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  CATGAGACTAATTGAGAACATCGGCGATCCTTCTCGGGCAGCATTTTCCACTCCAGAAGC  3180
```

```
Qy  3181  cggccgcgttagatccctgcaacagccacacccctccccgcctgcaatggatccagcttag  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  CGGCCGCCGTTAGATCCCTGCAACAGCCACACCTCCCCGTCTGCAATGGATCCAGCTTAG  3240

Qy  3241  cgccgagaggcgttctctacaatgtatagatctcactcacgcaccaccactctaaatga  3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  CGCCGAGAGAGCGGCGTTCTCTACAATGTATAGATCTCACACCTCACTCACGCCACCAACTCTAAATTGA  3300

Qy  3301  atcctaatttattaattcaccgatatctgaccgacccgttgaatttgtaggaagcga   3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  ATCCCTAATTATTAATTCACCGATATCTGACCGACCCGTTGAATTTTGTAGGAAGCGA  3360

Qy  3361  tcaggctgccgcactggaacatcaacgcacacagcatagtgtacgcgatcagaggacaag  3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  TCAGGCTGCCGCACTGGAACATCAACGCACACAGCATAGTGTACGCGATCAGAGGACAAG  3420

Qy  3421  ccagagtccagatcgtgaacgaggaagggaattcggtgttcgatggaggtgctgcaggag  3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  CCAGAGTCCAGATCGTGAACGAGGAAGGGAATTCGGTGTTCGATGGAGTGCTGCAGGAAG  3480

Qy  3481  gacaggtggtgacggtgccgcagaacttcgcggttgtaagagatcccagagcgagaggt  3540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  GACAGGTGGTGACGGTGCCGCAGAACTTCGCGGTTGTAAAGAGATCCCAGGAGCGAGAGGT  3540

Qy  3541  ttgatgggtggcgttcaaggacccagttcaagcgtgaactcgctagccgggagga      3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  TTGAGTGGGTGGCGTTCAAGACCCAGTTCAAGCGTGAACTCGCTAGCCGGGAGGA  3600

Qy  3601  catcgcagtaaggcgatcccgggatgttactggctaacgcctggaggtgtcgccgg     3660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  CATCGCAGTAAGGCGATCCCGCGGATGTTACTGGCTAACGCCTGGAGGTGTCGCCGG  3660

Qy  3661  aggagcggagagggtgaagttcaacagcaggacttcacttggctagdacaggggcc     3720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  AGGAGCGGCGAGGAGGTGAAGTTCAACAGCAGGAGACTCACTTGGCTAGCACCAGGGGCC  3720

Qy  3721  agtccagtcgcccgggagtgatgtctcaaggaggtgatcaactgcttctgctttgtaa    3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  AGTCCAGTCGCCCGGGAGGTGAATGTCGTCAAGGAGGTGATCAACTTGCTTATGTAAA  3780

Qy  3781  atgtgacggtgaaataacggtaaaataatatgtaataatataatataaagccaca    3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  ATGTGACGGTGAAATAATAACACGGTAAAATAATAACGGTAATAATAATAATAAAGCCACAA  3840

Qy  3841  agtgagaataaggggaaggggaaatgtgtaatgagccagtagccggtggtgctaattttg  3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  AGTGAGAATGGGGAGGGGAAATGTGTAATGTGTAATGAGCCAGTAGCCAGTAGCCCGTGGTCTAATTTTG  3900

Qy  3901  tatcgtattgtcaataaatcatgaattttgtgtttttatgtttttaaatcatgga    3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  TATCGTATTGTCAATAAATCATGAATTTTGTGGTTTTTAAATCATGAA  3960

Qy  3961  ttttaaattttataaaataatctccaatcggaagaacaacattccatatccatggatgtt  4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  TTTTAAATTTTATAAAATATCTCCAATCGGAAGAACAACAATTCCATATCCATGGATGTT  4020

Qy  4021  tcttttacccaaatctagttcttgagagatgaagcatcaccgaacagttctgcaactatc  4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  TCTTTACCCAAATCTAGTTCTTGAGAGGATGAAGCATCATCCATATCCGAACAGTTCTGCAACTATC  4080

Qy  4081  cctcaaagctttaaaatgaacaacaagagaacaacagagaacaacgttcaaaagatcccaaacga  4140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  CCTCAAAAGCTTTAAAATGAACAACAAGAGAACAAGAGCAACGTTCCAAAGATCCCAAACGA  4140

Qy  4141  aacatattatctatactaataatattaattactgccggaatcacaatccct            4200
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141  AACATATTATATCTATACTAATATTATTAATTACTACTCCCGGAATCACAATCACCCT  4200

Qy  4201  gaatgattcctattaactacagccttgttggcggcggaagtgatccgccggcggagaga  4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201  GAATGATTCCTATTAACTACACAAGCCTTGTTGGCGGCGCCTTGTTGGCGGCGGAGAGTGATCGGCGGCGGAGAGA  4260

Qy  4261  agcagcggactcggagacgaggcctggatgcagcagagtctttacctgccggcggtgaa  4320
```

```
Db  4261  ||||||||||||||||||||||||||||||||||||||||||||
          AGCAGCCGACTCGGAGACGAGGCCTTGGATGAGCAGAGTCTTACCTGCCAGGGCCTGAA 4320
Qy  4321  gggaagagcggccttttcggaagtaggagttcagcaagcggcggttccttgcggagtaag 4380
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4321  GGGGAAGAGCGGCCTTTCGGAGCATAGGAGTTCAGCAAGCGGCGGTTCCTTGGCGGAGTAAG 4380
Qy  4381  cggacgtaaggtggntgtcacgtcntcgttcnggagcgnattcatgaaggttaaa 4440
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4381  CGGACGTAAGGGTGGNNTGTCGACGTCNTCGTTTCNGGAGGCCGNATTCATGAAGGGTAAA 4440
Qy  4441  gtcanatctgagctctcgagtgctcaggggagccnaaaagacgttggaaccgtcgncgt 4500
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4441  GTCANATCTGAGCTCTCGAGTGCTCAGGGAGCCNAAAGACGTTGGGAAACCGTCGNCGT 4500
Qy  4501  ttgggcatcagtcngcgggcacgcttccctcctgctgtccanaancnangtanatt 4560
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4501  TTGGGCATCAGTCNGCGGGGCACGCTTCCCTCCTGCTGCTCCANAANCNANGTANATT 4560
Qy  4561  aaaaganatggqaaataataattaataggnaatnannaggaggattgnaacggtcnganccgn 4620
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4561  AAAAGANATGGGAAATTAANTAATGGNAATNANNAGGAGGATTGNAACGGTCNGANCCGN 4620
Qy  4621  angaaagttttannggtttaaatactgggagtgngagccngccnctgttccngtg 4680
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4621  ANGAAANAGTTTTANNGGTTTAAATACTGGGGAGTGNAGCCNGGCNGCCNCTGGTTCCNGTG 4680
Qy  4681  taqangaaaccaagnnnccgggagyntncannngnaggggagaaaaggganncatttnan 4740
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4681  TAGANGAAACCAAGNNCCGGGAGGNTTNCANNNGNNNAGGGAGAAAAAGGANNCATTTNAN 4740
Qy  4741  nangcnaggaggacatgaancggtacnagagctgnggttcannnancggcgnnnggnagtcc 4800
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4741  NANGCNGAGGGACCATGAANCGGTACNGAGCTGNGGTTCANNNANCGGCGNNNNGGNAGTCC 4800
Qy  4801  cnnggaccngngtgggtnanaagggaangggaacattngngnnganggaaanaccntt 4860
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4801  CNNGGACCNGNGTGGGGTNANAAGGGAANGGGAANCANTTNGGTNGNANGGANAANACCNTT 4860
Qy  4861  ttacnattgccttgcaagnnngtntnggcncntncnggtnacatnccgctcgatggct 4920
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4861  TTACNATTGCCTTGCAGGNNGTNTNGGCNCNTNCNGGTNACATNCCGCTCGCATGGGCT 4920
Qy  4921  ttggggngcnanaggnagccncangggnanncngccncctgtncangncgctnaagtt 4980
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTGGGGNCCCNANAGGNAGCCNCANAGGGNANNCNGCCNCCTTGTNCANGNCGCTNAAGTT 4980
Qy  4981  cnattgtanatggngcgttg 4999
          ||||||||||||||||||||
Db  4981  CNATTGTANATGGNGCGTTG 4999
```

RESULT   2
ACO95913
LOCUS   ACO95913   26541 bp   DNA   linear   HTG   20-DEC-2001
DEFINITION   Rattus norvegicus clone CH230-10M9, *** SEQUENCING IN PROGRESS ***,
             17 unordered pieces.
ACCESSION   ACO95913
VERSION   ACO95913.3   GI:17943549
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Norway rat.
ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE   1 (bases 1 to 26541)
AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
             Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
             Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
             Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
             Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
             Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
             Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.C.,
             Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
             Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
             Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
             Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
             Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
             Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
             Garza,A., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
             Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
             Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
             Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
             Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
             Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
             Kovar,C., Kratovic,J., Kureshi,A., Kelly,N., Leal,B., Lewis,L.C.,
             Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
             Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
             Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
             Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
             Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
             Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newtson,N.,
             Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
             Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
             Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
             Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
             Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
             Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
             Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
             Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
             Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
             Wall,R., Wang,S., Ward-Moore,S., Warren,A., Washington,C.,
             Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
             Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
             Weinstock,G. and Gibbs,R.
TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 26541)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
             of Molecular and Human Genetics, Baylor College of Medicine, One
             Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Dec 20, 2001 this sequence version replaced gi:16901678.
             ------------------ Genome Center
             Center: Baylor College of Medicine
             Center code: BCM
             Web site: http://www.hgsc.bcm.tmc.edu/
             Contact: hgsc-help@bcm.tmc.edu
             ------------------ Project Information
             Center project name: GDVO
             Center clone name: CH230-10M9
             ------------------ Summary Statistics
             Assembly program: Phrap; version 0.990329First call to
             findPhraplist
             Consensus quality: 13487 bases at least Q40
             Consensus quality: 15808 bases at least Q30
             Consensus quality: 17668 bases at least Q20
             Estimated insert size: 10184; sum-of-contigs estimation
             Quality coverage: 0x in Q20 bases: agarose-fp estimation
             Quality coverage: 0.1x in Q20 bases; sum-of-contigs estimation
             * NOTE: Estimated insert size may differ from sequence length
             *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 17 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *   1   1829: contig of 1829 bp in length
             * 1830   1929: gap of unknown length
             * 1930   3748: contig of 1819 bp in length
             * 3749   3848: gap of unknown length
```

```
     *        3849    5491: contig of 1643 bp in length
     *        5492    5591: gap of unknown length
     *        5592    6820: contig of 1229 bp in length
     *        6821    6920: gap of unknown length
     *        6921    9295: contig of 2375 bp in length
     *        9296    9395: gap of unknown length
     *        9396   10779: contig of 1384 bp in length
     *       10780   10879: gap of unknown length
     *       10880   11955: contig of 1076 bp in length
     *       11956   12055: gap of unknown length
     *       12056   13528: contig of 1473 bp in length
     *       13529   13628: gap of unknown length
     *       13629   14870: contig of 1242 bp in length
     *       14871   14970: gap of unknown length
     *       14971   16577: contig of 1607 bp in length
     *       16578   16677: gap of unknown length
     *       16678   17927: contig of 1250 bp in length
     *       17928   18027: gap of unknown length
     *       18028   19217: contig of 1190 bp in length
     *       19218   19317: gap of unknown length
     *       19318   20918: contig of 1601 bp in length
     *       20919   21018: gap of unknown length
     *       21019   22505: contig of 1487 bp in length
     *       22506   22605: gap of unknown length
     *       22606   24325: contig of 1720 bp in length
     *       24326   24425: gap of unknown length
     *       24426   25429: contig of 1004 bp in length
     *       25430   25529: gap of unknown length
     *       25530   26541: contig of 1012 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..26541
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-10M9"
BASE COUNT     6642 a   5472 c   5923 g   6839 t   1665 others
ORIGIN
```

```
Query Match             0.5%;  Score 25;  DB 2;  Length 26541;
Best Local Similarity 100.0%;  Pred. No. 0.11;
Matches  25;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    3808 atatatgtaataataataataa 3832
           ||||||||||||||||||||||
Db   16191 ATATATGTAATAATAATAATAA 16215
```

```
RESULT  3
AC108615
LOCUS       AC108615          113725 bp    DNA     linear   HTG 31-JAN-2002
DEFINITION  Rattus norvegicus clone CH230-290E9, *** SEQUENCING IN PROGRESS
            ***, 62 unordered pieces.
ACCESSION   AC108615
VERSION     AC108615.1  GI:18449450
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 113725)
AUTHORS     Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,U., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchel,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newton,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,F., Primus,E., Pu,L.L., Qulles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 113725)
AUTHORS     Worley,K.C.
TITLE       Direct Submission
JOURNAL     Submitted (31-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            -------------------- Project Information
            Center project name: GPOT
            Center clone name: CH230-290E9
            -------------------- Summary Statistics
            Sequencing vector: Plasmid: M77789
            Chemistry: Dye-terminator Big Dye: 100% of reads
            Assembly program: Phrap: version 0.990329First call to
            findPhrapList
            Consensus quality: 87886 bases at least Q40
            Consensus quality: 94485 bases at least Q30
            Consensus quality: 99429 bases at least Q20
            Estimated insert size: 82890; sum-of-contigs estimation
            Quality coverage: 0x in Q20 bases; agarose-fp estimation
            Quality coverage: 1x in Q20 bases; sum-of-contigs estimation
            * NOTE: Estimated insert size may differ from sequence length
            *       (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            *       consists of 62 contigs. The true order of the pieces
            *       is not known and their order in this sequence record is
            *       arbitrary. Gaps between the contigs are represented as
            *       runs of N, but the exact sizes of the gaps are unknown.
            *       This record will be updated with the finished sequence
            *       as soon as it is available and the accession number will
            *       be preserved.
            *         1    3973: contig of 3973 bp in length
            *      3974    4073: gap of unknown length
            *      4074    7436: contig of 3363 bp in length
            *      7437    7536: gap of unknown length
            *      7537   10805: contig of 3269 bp in length
            *     10806   10905: gap of unknown length
            *     10906   13707: contig of 2802 bp in length
```

```
*  13708  13807: gap of unknown length
*  13808  16732: contig of 2925 bp in length
*  16733  16832: gap of unknown length
*  16833  18977: contig of 2145 bp in length
*  18978  19077: gap of unknown length
*  19078  21825: contig of 2748 bp in length
*  21826  21925: gap of unknown length
*  21926  24091: contig of 2166 bp in length
*  24092  24191: gap of unknown length
*  24192  25964: contig of 1773 bp in length
*  25965  26064: gap of unknown length
*  26065  28909: contig of 2845 bp in length
*  28910  29009: gap of unknown length
*  29010  30413: contig of 1404 bp in length
*  30414  30513: gap of unknown length
*  30514  31753: contig of 1240 bp in length
*  31754  31853: gap of unknown length
*  31854  34448: contig of 2595 bp in length
*  34449  34548: gap of unknown length
*  34549  36351: contig of 1803 bp in length
*  36352  36451: gap of unknown length
*  36452  38532: contig of 2081 bp in length
*  38533  38632: gap of unknown length
*  38633  40773: contig of 2141 bp in length
*  40774  40873: gap of unknown length
*  40874  42725: contig of 1852 bp in length
*  42726  42825: gap of unknown length
*  42826  44501: contig of 1676 bp in length
*  44502  44601: gap of unknown length
*  44602  46744: contig of 2143 bp in length
*  46745  46844: gap of unknown length
*  46845  48727: contig of 1883 bp in length
*  48728  48827: gap of unknown length
*  48828  51179: contig of 2352 bp in length
*  51180  51279: gap of unknown length
*  51280  52478: contig of 1199 bp in length
*  52479  52578: gap of unknown length
*  52579  54000: contig of 1422 bp in length
*  54001  54100: gap of unknown length
*  54101  55860: contig of 1760 bp in length
*  55861  55960: gap of unknown length
*  55961  58011: contig of 2051 bp in length
*  58012  58111: gap of unknown length
*  58112  59783: contig of 1672 bp in length
*  59784  59883: gap of unknown length
*  59884  61804: contig of 1921 bp in length
*  61805  61904: gap of unknown length
*  61905  63014: contig of 1110 bp in length
*  63015  63114: gap of unknown length
*  63115  64428: contig of 1314 bp in length
*  64429  64528: gap of unknown length
*  64529  65848: contig of 1320 bp in length
*  65849  65948: gap of unknown length
*  65949  67951: contig of 2003 bp in length
*  67952  68051: gap of unknown length
*  68052  69813: contig of 1762 bp in length
*  69814  69913: gap of unknown length
*  69914  71245: contig of 1332 bp in length
*  71246  71345: gap of unknown length
*  71346  72906: contig of 1561 bp in length
*  72907  73006: gap of unknown length
*  73007  74675: contig of 1669 bp in length
*  74676  74775: gap of unknown length
*  74776  76188: contig of 1413 bp in length
*  76189  76288: gap of unknown length
*  76289  77815: contig of 1527 bp in length
*  77816  77915: gap of unknown length
*  77916  79872: contig of 1957 bp in length
*  79873  79972: gap of unknown length
*  79973  81380: contig of 1408 bp in length
*  81381  81480: gap of unknown length
*  81481  82512: contig of 1032 bp in length
*  82513  82612: gap of unknown length
*  82613  84502: contig of 1890 bp in length
*  84503  84602: gap of unknown length
*  84603  85955: contig of 1353 bp in length
*  85956  86055: gap of unknown length
*  86056  87130: contig of 1075 bp in length
*  87131  87230: gap of unknown length
*  87231  88779: contig of 1549 bp in length
*  88780  88879: gap of unknown length
*  88880  90359: contig of 1480 bp in length
*  90360  90459: gap of unknown length
*  90460  91498: contig of 1039 bp in length
*  91499  91598: gap of unknown length
*  91599  93463: contig of 1865 bp in length
*  93464  93563: gap of unknown length
*  93564  94976: contig of 1413 bp in length
*  94977  95076: gap of unknown length
*  95077  96234: contig of 1158 bp in length
*  96235  96334: gap of unknown length
*  96335  97663: contig of 1329 bp in length
*  97664  97763: gap of unknown length
*  97764  99250: contig of 1487 bp in length
*  99251  99350: gap of unknown length
*  99351 100809: contig of 1459 bp in length
* 100810 100909: gap of unknown length
* 100910 101932: contig of 1023 bp in length
* 101933 102032: gap of unknown length
* 102033 103042: contig of 1010 bp in length
* 103043 103142: gap of unknown length
* 103143 104157: contig of 1015 bp in length
* 104158 104257: gap of unknown length
```

```
Query Match       0.5%;   Score 25;   DB 2;   Length 113725;
Best Local Similarity   100.0%;  Pred. No. 0.11;
Matches   25;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   3808 atatatgtaataataataataa 3832
          |||||||||||||||||||||||||
Db  13816 ATATATGTAATAATAATAATAATAA 13840
```

```
RESULT 4
AL136233/c
LOCUS       AL136233        138733 bp    DNA      linear   PRI 21-OCT-2001
DEFINITION  Human DNA sequence from clone RP11-9E13 on chromosome 10, complete
            sequence.
ACCESSION   AL136233
VERSION     AL136233.14   GI:16408601
KEYWORDS    HTG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 138733)
AUTHORS     Howden,P.
TITLE       Direct Submission
JOURNAL     Submitted (20-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 24, 2001 this sequence version replaced gi:15131189.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e. phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
```

```
                                                             /note="Primer"
BASE COUNT        12 a       4 c       9 g       7 t
ORIGIN

Query Match
Best Local Similarity   0.5%;   Score 24;  DB 6;  Length 32;
Matches   24;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   3780 aatgtgacggtgaaataataacgg 3803
          |||||||||||||||||||||||
Db      9 AATGTGACGGTGAAATAATAACGG 32


RESULT 6
AF240004
LOCUS       AF240004                1683 bp    mRNA    linear    PLN 02-MAR-2001
DEFINITION  Sesamum indicum 11s globulin mRNA, complete cds.
ACCESSION   AF240004
VERSION     AF240004.1  GI:13183172
KEYWORDS    .
SOURCE      Sesamum indicum
ORGANISM    Sesamum indicum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.
REFERENCE   1 (bases 1 to 1683)
  AUTHORS   Tai,S.S.K. and Tzen,J.T.C.
  TITLE     Molecular cloning of three storage proteins in sesame
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1683)
  AUTHORS   Tai,S.S.K. and Tzen,J.T.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2000) Graduate Institute of Agricultural
            Biotechnology, National Chung-Hsing University, Taichung 40227,
            Taiwan
FEATURES             Location/Qualifiers
     source          1..1683
                     /organism="Sesamum indicum"
                     /db_xref="taxon:4182"
     CDS             24..1517
                     /note="seed storage protein; storage protein in seed
                     protein bodies"
                     /codon_start=1
                     /product="11S globulin"
                     /protein_id="AAK15087.1"
                     /db_xref="GI:13183173"
                     /translation="MALTSLLSFFIVTLLIRGLSAQLAGEQDFYWQDLQSQQHKLQ
                     ARTDCRVERLTAQEPIRFESEAGLTEFWDRNNQQFECAGVAAVRNVIQPRGLLLPHY
                     NNAPQLLYVVRGRGIQGTVIPGCAETFERDTQPRQDRRRRFMDRHQKVRQFRQGDILA
                     LPAGLTLWFYNNGEPLITVALLDTGNAANQLDQTFRHFFLAGNPQGGRQSYFGRPQT
                     EKQQGETKNIFNGFDDEILADAFGVDVQTARRLKQODDLRGRIVRAERLDIVLPGEEE
                     EERWERDPYSGANGLEETLCTAKLRENLDEPARADVYNPHGGRISSLNSLTLPVLSWL
                     RLSAEKGVLYRNGLVAPHWNLNAHSIIYITRGSGRFQVVGHTGRSVFDGVVREQLII
                     VPQNYVVAKRASQDEGLEWISFKTNDNAMTSQLAGRLSAIRAMPEVVWTAYQVSRDE
                     ARRLKYNREESRVFSSTSRYSWPRSSRPMSYMPKFEYVLDVIKSMM"
BASE COUNT       442 a     397 c     477 g     367 t
ORIGIN

Query Match
Best Local Similarity   0.5%;   Score 24;  DB 8;  Length 1683;
Matches   24;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   3555 ttcaagaccaacgacaacgcgatg 3578
          ||||||||||||||||||||||||
Db   1263 TTCAAGACCAACGACAACGCGGATG 1286


RESULT 7
AC034196
LOCUS       AC034196              133108 bp    DNA     linear    HTG 29-MAY-2000
DEFINITION  Homo sapiens chromosome 3 clone RP11-710E4 map 3p, WORKING DRAFT
```

```
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-9E13 is from the library RPCI-11.1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-9E13 It may be shorter because we sequence overlapping
            sections only once, except for a short overlap.
            The true right end of clone RP11-9E13 is at 138733 in this
            sequence. The true right end of clone RP11-153K11 is at 2000 in
            this sequence.
FEATURES             Location/Qualifiers
     source          1..138733
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-9E13"
                     /clone_lib="RPCI-11.1"
     misc_feature    55963..56014
                     /note="Sequence from uni-directional dGTP big dye
                     terminator reads only."
     misc_feature    62420..63555
                     /note="CpG island"
                     /evidence=not_experimental
     misc_feature    67805..67836
                     /note="Single clone region. Assembly confirmed by
                     restriction digest data."
     misc_feature    77610..77625
                     /note="Single clone region. Assembly confirmed by
                     restriction digest data."
     misc_feature    127610..128010
                     /note="CpG island"
                     /evidence=not_experimental
BASE COUNT     43280 a   29780 c   28584 g   37089 t
ORIGIN

Query Match
Best Local Similarity   0.5%;   Score 25;  DB 9;  Length 138733;
Matches   25;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   3815 taataataataataaagccaca 3839
          ||||||||||||||||||||||
Db 108555 TAATAATAATAATAAAGCCACA 108531


RESULT 5
AX089372
LOCUS       AX089372                  32 bp    DNA     linear    PAT 21-MAR-2001
DEFINITION  Sequence 22 from Patent WO0116340.
ACCESSION   AX089372
VERSION     AX089372.1  GI:13443635
KEYWORDS    .
SOURCE      synthetic construct.
ORGANISM    synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 32)
  AUTHORS   Chaudhary,S., van Rooijen,G., Moloney,M.M. and Singh,S.
  TITLE     Flax seed specific promoters
  JOURNAL   Patent: WO 0116340-A 22 08-MAR-2001;
            SemBioSys Genetics Inc. (CA) ; COMMONWEALTH SCIENTIFIC AND
            INDUSTRIAL RESEARCH ORGANISATION (AU)
FEATURES             Location/Qualifiers
     source          1..32
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
```

```
          SEQUENCE, 38 unordered pieces.
ACCESSION   AC034196
VERSION     AC034196.2  GI:8101269
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 133108)
  AUTHORS   Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
            Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
            Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,G., Li,J., Li,L.,
            Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,Y., Li,Y.,
            Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
            Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
            Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
            Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
            Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
            Zhang,Z., Zhu,B., Yu,J. and Yang,H.
  TITLE     Chromosome 3p genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 133108)
  AUTHORS   Li,W., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B., Bao,W.,
            Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y., Qi,X.,
            Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Li,G.,
            Li,C., Bao,Q., Bao,J., Wang,J., Song,L., Guo,D.,
            Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
            and Yang,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-APR-2000) Human Genomic Center, Institute of
            Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
            100101,P.R.China
COMMENT     On May 29, 2000 this sequence version replaced gi:7417469.
            ----------Genome Center
            Center:Beijing Center
            Center code:Beijing
            Website:http://hgc.igtp.ac.cn
            http://www.genomics.org.cn
            Contact:hgc@igtp.ac.cn
            ---------- Project Information
            Center project name:1% project
            Center clone name: RP11-710E4
            ---------- Summary Statistics
            Sequencing vector: pUC18; 100% of reads
            Chemistry: Dye-terminator: ET 55% of reads
            Chemistry: Dye-terminator Big Dye: 45% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 110057 bases at least Q40
            Consensus quality: 124151 bases at least Q30
            Consensus quality: 133102 bases at least Q20
            Insert size: 102269; sum-of-contigs
            Quality coverage:   3.16x in Q20 bases;sum-of-contigs
            ----------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 38 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      1163: contig of 1163 bp in length
            *     1164      1263: gap of unknown length
            *     1264      2750: contig of 1487 bp in length
            *     2751      2850: gap of unknown length
            *     2851      3964: contig of 1114 bp in length
            *     3965      4064: gap of unknown length
            *     4065      5164: contig of 1100 bp in length
            *     5165      5264: gap of unknown length
            *     5265      6462: contig of 1198 bp in length
            *     6463      6562: gap of unknown length
            *     6563      7875: contig of 1313 bp in length
            *     7876      7975: gap of unknown length
            *     7976      9431: contig of 1456 bp in length
            *     9432      9531: gap of unknown length
            *     9532     11077: contig of 1546 bp in length
            *    11078     11177: gap of unknown length
            *    11178     12990: contig of 1813 bp in length
            *    12991     13090: gap of unknown length
            *    13091     15234: contig of 2144 bp in length
            *    15235     15334: gap of unknown length
            *    15335     17069: contig of 1735 bp in length
            *    17070     17169: gap of unknown length
            *    17170     18806: contig of 1637 bp in length
            *    18807     18906: gap of unknown length
            *    18907     20498: contig of 1592 bp in length
            *    20499     20598: gap of unknown length
            *    20599     23090: contig of 2492 bp in length
            *    23091     23190: gap of unknown length
            *    23191     26302: contig of 3112 bp in length
            *    26303     26402: gap of unknown length
            *    26403     28833: contig of 2431 bp in length
            *    28834     28933: gap of unknown length
            *    28934     31329: contig of 2396 bp in length
            *    31330     31429: gap of unknown length
            *    31430     33222: contig of 1793 bp in length
            *    33223     33322: gap of unknown length
            *    33323     35570: contig of 2248 bp in length
            *    35571     35670: gap of unknown length
            *    35671     37854: contig of 2184 bp in length
            *    37855     37954: gap of unknown length
            *    37955     39204: contig of 1250 bp in length
            *    39205     39304: gap of unknown length
            *    39305     42732: contig of 3428 bp in length
            *    42733     42832: gap of unknown length
            *    42833     45249: contig of 2417 bp in length
            *    45250     45349: gap of unknown length
            *    45350     48684: contig of 3335 bp in length
            *    48685     48784: gap of unknown length
            *    48785     52145: contig of 3361 bp in length
            *    52146     52245: gap of unknown length
            *    52246     55937: contig of 3692 bp in length
            *    55938     56037: gap of unknown length
            *    56038     59494: contig of 3457 bp in length
            *    59495     59594: gap of unknown length
            *    59595     64439: contig of 4845 bp in length
            *    64440     64539: gap of unknown length
            *    64540     69590: contig of 5051 bp in length
            *    69591     69690: gap of unknown length
            *    69691     74741: contig of 5051 bp in length
            *    74742     74841: gap of unknown length
            *    74842     79798: contig of 4957 bp in length
            *    79799     79898: gap of unknown length
            *    79899     84721: contig of 4823 bp in length
            *    84722     84821: gap of unknown length
            *    84822     89494: contig of 4673 bp in length
            *    89495     89594: gap of unknown length
            *    89595     96167: contig of 6573 bp in length
            *    96168     96267: gap of unknown length
            *    96268    105799: contig of 9532 bp in length
            *   105800    105899: gap of unknown length
            *   105900    114293: contig of 8394 bp in length
            *   114294    114393: gap of unknown length
            *   114394    123481: contig of 9088 bp in length
            *   123482    123581: gap of unknown length
            *   123582    133108: contig of 9527 bp in length.
FEATURES             Location/Qualifiers
     source          1..133108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /map="3p"
                     /clone="RP11-710E4"
     misc_feature    1..1163
                     /note="assembly_name:Contig12"
     misc_feature    1264..2750
```

```
     misc_feature     2851..3964
                       /note="assembly_name:Contig16"
     misc_feature     4065..5164
                       /note="assembly_name:Contig17"
     misc_feature     5265..6462
                       /note="assembly_name:Contig20"
     misc_feature     6563..7875
                       /note="assembly_name:Contig23"
     misc_feature     7976..9431
                       /note="assembly_name:Contig26"
                       /note="assembly_name:Contig28"
     misc_feature     9532..11077
                       /note="assembly_name:Contig29"
     misc_feature     11178..12990
                       /note="assembly_name:Contig30
                       clone_end:T7
                       vector_side:right"
     misc_feature     13091..15234
                       /note="assembly_name:Contig32"
     misc_feature     15335..17069
                       /note="assembly_name:Contig33"
     misc_feature     17170..18806
                       /note="assembly_name:Contig34"
     misc_feature     18907..20498
                       /note="assembly_name:Contig35"
     misc_feature     20599..23090
                       /note="assembly_name:Contig36"
     misc_feature     23191..26302
                       /note="assembly_name:Contig37"
     misc_feature     26403..28833
                       /note="assembly_name:Contig38"
     misc_feature     28934..31329
                       /note="assembly_name:Contig39"
     misc_feature     31430..33222
                       /note="assembly_name:Contig40"
     misc_feature     33323..35570
                       /note="assembly_name:Contig41"
     misc_feature     35671..37854
                       /note="assembly_name:Contig42"
     misc_feature     37955..39204
                       /note="assembly_name:Contig43"
     misc_feature     39305..42732
                       /note="assembly_name:Contig44"
     misc_feature     42833..45249
                       /note="assembly_name:Contig45"
     misc_feature     45350..48684
                       /note="assembly_name:Contig46"
     misc_feature     48785..52145
                       /note="assembly_name:Contig47"
     misc_feature     52246..55937
                       /note="assembly_name:Contig48"
     misc_feature     56038..59494
```

```
Query Match     0.5%;  Score 24;  DB 2;  Length 133108;
Best Local Similarity   100.0%;  Pred. No. 0.4;
Matches  24;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    3803 gtaaatatgtaataataaa 3826
           |||||||||||||||||||||
Db  120755 GTAAAATATGTAATAATAATAA 120778
```

```
RESULT    8
AC022096
LOCUS       AC022096          148598 bp    DNA     linear    HTG 18-JUL-2000
DEFINITION  Homo sapiens chromosome 5 clone CTB-43D14, WORKING DRAFT SEQUENCE,
            20 ordered pieces.
ACCESSION   AC022096
VERSION     AC022096.4  GI:9256401
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 148598)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 148598)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 18, 2000 this sequence version replaced gi:7711677.
            -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----
            Project Information
            Center Project Name: 80802
            Center clone name: CIT978SKB_43D14
            -----
            Summary Statistics
            Consensus quality: 133423 bases at least Q40
            Consensus quality: 144004 bases at least Q30
            Consensus quality: 145802 bases at least Q20
            Estimated insert size: 176000; pulse field gel estimation
            Estimated insert size: 147698; sum-of-contigs estimation
            Quality coverage: 5.34 in Q20 bases; pulse field gel estimation
            Quality coverage: 6.36 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 20 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *          1     5195: contig of 5195 bp in length
            *       5196     5295: gap of unknown length
            *       5296    15425: contig of 10130 bp in length
            *      15426    15525: gap of unknown length
            *      15526    20947: contig of 5422 bp in length
            *      20948    21047: gap of unknown length
            *      21048    26789: contig of 5742 bp in length
            *      26790    26889: gap of unknown length
            *      26890    32943: contig of 6054 bp in length
            *      32944    33043: gap of unknown length
            *      33044    38651: contig of 5608 bp in length
            *      38652    38751: gap of unknown length
            *      38752    64412: contig of 25661 bp in length
            *      64413    64512: gap of unknown length
            *      64513    79245: contig of 14733 bp in length
            *      79246    79345: gap of unknown length
            *      79346    82132: contig of 2787 bp in length
            *      82133    82232: gap of unknown length
            *      82233    86478: contig of 4246 bp in length
            *      86479    86578: gap of unknown length
            *      86579    90700: contig of 4122 bp in length
            *      90701    90800: gap of unknown length
            *      90801    93106: contig of 2306 bp in length
            *      93107    93206: gap of unknown length
            *      93207    99356: contig of 6150 bp in length
            *      99357    99456: gap of unknown length
            *      99457   116776: contig of 17320 bp in length
            *     116777   116876: gap of unknown length
            *     116877   117926: contig of 1050 bp in length
            *     117927   118026: gap of unknown length
            *     118027   127710: contig of 9684 bp in length
            *     127711   127810: gap of unknown length
            *     127811   136755: contig of 8945 bp in length
            *     136756   136855: gap of unknown length
            *     136856   138975: contig of 2120 bp in length
            *     138976   139075: gap of unknown length
```

```
*       139076    146683: contig of 7608 bp in length
*       146684    146783: gap of unknown length
*       146784    148598: contig of 1815 bp in length.
FEATURES             Location/Qualifiers
     source          1..148598
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTB-43D14"
                     /clone_lib="Caltech human BAC library B"
BASE COUNT    41000 a  35188 c  33298 g  37210 t   1902 others
ORIGIN

Query Match
Best Local Similarity  100.0%;  Score 24;  DB 2;  Length 148598;
Matches  24;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   3955 catgaattttataaaat 3978
          |||||||||||||||||
Db 124720 CATGAATTTTAAATTTATAAAT 124743


RESULT  9
AC026193/c
LOCUS       AC026193              152602 bp    DNA     linear   HTG 29-MAY-2000
DEFINITION  Homo sapiens chromosome 3 clone RP11-416H8 map 3p, WORKING DRAFT
            SEQUENCE, 57 unordered pieces.
ACCESSION   AC026193
VERSION     AC026193.2  GI:8101213
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152602)
  AUTHORS   Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
            Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
            Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,G., Li,J., Li,L.,
            Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y.,
            Luo,J., Niu,Y., Qi,X., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
            Tao,R., Wang,H., Wang,J., Wang,J., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
            Wang,X., Wang,Y., Wang,Y., Yu,D., Yang,G., Zhang,H., Zhang,H.,
            Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
            Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
            Zhang,Z., Zhu,B., Yu,J. and Yang,H.
  TITLE     Chromosome 3p genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152602)
  AUTHORS   Sun,W., Hu,S., Dong,W., Wang,J., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
            Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
            Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Guo,D.,
            Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,M., Li,L., Guo,D.,
            Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
            and Yang,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Human Genomic Center, Institute of
            Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
            100101, P.R.China
COMMENT     On May 29, 2000 this sequence version replaced gi:7272010.
            ----------------Genome Center
            Center:Beijing Center
            Center code:Beijing
            Website:http://hgc.igtp.ac.cn
            http://www.genomics.org.cn
            Contact:hgc@igtp.ac.cn
            ---------------- Project Information
            Center project name:1% project
            Center clone name: RP11-416H8
            ---------------- Summary Statistics
            Sequencing vector: pUC18; 100% of reads
            Chemistry: Dye-terminator: ET 55% of reads
            Chemistry: Dye-terminator Big Dye: 45% of reads
```

```
Assembly program: Phrap; version 0.990329
Consensus quality: 115486 bases at least Q40
Consensus quality: 144075 bases at least Q30
Consensus quality: 163107 bases at least Q20
Insert size: 88668; sum-of-contigs
Quality coverage:  3.78x in Q20 bases;sum-of-contigs
--------
* NOTE: This is a 'working draft' sequence. It currently
* consists of 57 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
*        1    1540:  contig of 1540 bp in length
*     1541    1640:  gap of unknown length
*     1641    2835:  contig of 1195 bp in length
*     2836    2935:  gap of unknown length
*     2936    4034:  contig of 1099 bp in length
*     4035    4134:  gap of unknown length
*     4135    5760:  contig of 1626 bp in length
*     5761    5860:  gap of unknown length
*     5861    7021:  contig of 1161 bp in length
*     7022    7121:  gap of unknown length
*     7122    8431:  contig of 1310 bp in length
*     8432    8531:  gap of unknown length
*     8532   10150:  contig of 1619 bp in length
*    10151   10250:  gap of unknown length
*    10251   11827:  contig of 1577 bp in length
*    11828   11927:  gap of unknown length
*    11928   13240:  contig of 1313 bp in length
*    13241   13340:  gap of unknown length
*    13341   14802:  contig of 1462 bp in length
*    14803   14902:  gap of unknown length
*    14903   16661:  contig of 1759 bp in length
*    16662   16761:  gap of unknown length
*    16762   18070:  contig of 1309 bp in length
*    18071   18170:  gap of unknown length
*    18171   19601:  contig of 1431 bp in length
*    19602   19701:  gap of unknown length
*    19702   21364:  contig of 1663 bp in length
*    21365   21464:  gap of unknown length
*    21465   23376:  contig of 1912 bp in length
*    23377   23476:  gap of unknown length
*    23477   25479:  contig of 2003 bp in length
*    25480   25579:  gap of unknown length
*    25580   27646:  contig of 2067 bp in length
*    27647   27746:  gap of unknown length
*    27747   29548:  contig of 1802 bp in length
*    29549   29648:  gap of unknown length
*    29649   31701:  contig of 2053 bp in length
*    31702   31801:  gap of unknown length
*    31802   33144:  contig of 1343 bp in length
*    33145   33244:  gap of unknown length
*    33245   34535:  contig of 1291 bp in length
*    34536   34635:  gap of unknown length
*    34636   35728:  contig of 1093 bp in length
*    35729   35828:  gap of unknown length
*    35829   37639:  contig of 1811 bp in length
*    37640   37739:  gap of unknown length
*    37740   40142:  contig of 2403 bp in length
*    40143   40242:  gap of unknown length
*    40243   41570:  contig of 1328 bp in length
*    41571   41670:  gap of unknown length
*    41671   43259:  contig of 1589 bp in length
*    43260   43359:  gap of unknown length
*    43360   45147:  contig of 1788 bp in length
*    45148   45247:  gap of unknown length
*    45248   46975:  contig of 1728 bp in length
*    46976   47075:  gap of unknown length
*    47076   49136:  contig of 2061 bp in length
*    49137   49236:  gap of unknown length
```

```
misc_feature      /note="assembly_name:Contig57"
                  8532..10150
misc_feature      /note="assembly_name:Contig62"
                  10251..11827
misc_feature      /note="assembly_name:Contig63"
                  11928..13240

Query Match              0.5%;   Score 24;  DB 2;  Length 152602;
Best Local Similarity  100.0%;  Pred. No. 0.4;
Matches   24;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    3803 gtaaaatatgtaataataa 3826
           |||||||||||||||||||
Db 148305 GTAAAATATGTAATAATAA 148282

RESULT 10
AC024377
LOCUS       AC024377              162433 bp    DNA      linear   HTG 26-MAY-2000
DEFINITION  Homo sapiens clone RP11-611H22, WORKING DRAFT SEQUENCE, 24
            unordered pieces.
ACCESSION   AC024377
VERSION     AC024377.3  GI:8076825
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 162433)
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens, clone RP11-611H22
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 162433)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Diaz,J.S., Dodge,M., Domino,M., Doyle,M.,
            Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., Landers,T., Largocque,K., Lehoczky,J., Levine,R.,
            Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
            McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M.,
            Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
            Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
            Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
            Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 25, 2000 this sequence version replaced gi:7321591.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www.seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
            Center project name: L6258
            Center clone name: 611_H_22
            -------------- Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
```

```
*    49237    52178: contig of 2942 bp in length
*    52179    52278: gap of unknown length
*    52279    54957: contig of 2679 bp in length
*    54958    55057: gap of unknown length
*    55058    57417: contig of 2360 bp in length
*    57418    57517: gap of unknown length
*    57518    60195: contig of 2678 bp in length
*    60196    60295: gap of unknown length
*    60296    62295: contig of 2000 bp in length
*    62296    62395: gap of unknown length
*    62396    64401: contig of 2006 bp in length
*    64402    64501: gap of unknown length
*    64502    67886: contig of 3385 bp in length
*    67887    67986: gap of unknown length
*    67987    70888: contig of 2902 bp in length
*    70889    70988: gap of unknown length
*    70989    73943: contig of 2955 bp in length
*    73944    74043: gap of unknown length
*    74044    76750: contig of 2707 bp in length
*    76751    76850: gap of unknown length
*    76851    79324: contig of 2474 bp in length
*    79325    79424: gap of unknown length
*    79425    81198: contig of 1774 bp in length
*    81199    81298: gap of unknown length
*    81299    84892: contig of 3594 bp in length
*    84893    84992: gap of unknown length
*    84993    87653: contig of 2661 bp in length
*    87654    87753: gap of unknown length
*    87754    91151: contig of 3398 bp in length
*    91152    91251: gap of unknown length
*    91252    93951: contig of 2700 bp in length
*    93952    94051: gap of unknown length
*    94052    98226: contig of 4175 bp in length
*    98227    98326: gap of unknown length
*    98327   102109: contig of 3783 bp in length
*   102110   102209: gap of unknown length
*   102210   105933: contig of 3724 bp in length
*   105934   106033: gap of unknown length
*   106034   109844: contig of 3811 bp in length
*   109845   109944: gap of unknown length
*   109945   112749: contig of 2805 bp in length
*   112750   112849: gap of unknown length
*   112850   117845: contig of 4996 bp in length
*   117846   117945: gap of unknown length
*   117946   122444: contig of 4499 bp in length
*   122445   122544: gap of unknown length
*   122545   128723: contig of 6179 bp in length
*   128724   128823: gap of unknown length
*   128824   134603: contig of 5780 bp in length
*   134604   134703: gap of unknown length
*   134704   140178: contig of 5475 bp in length
*   140179   140278: gap of unknown length
*   140279   147021: contig of 6743 bp in length
*   147022   147121: gap of unknown length
*   147122   152602: contig of 5481 bp in length.

FEATURES          Location/Qualifiers
source            1..152602
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="3"
                  /map="3p"
                  /clone="RP11-416H8"
misc_feature      1..1540
                  /note="assembly_name:Contig42"
misc_feature      1641..2835
                  /note="assembly_name:Contig45"
misc_feature      2936..4034
                  /note="assembly_name:Contig46"
misc_feature      4135..5760
                  /note="assembly_name:Contig51"
misc_feature      5861..7021
                  /note="assembly_name:Contig52"
misc_feature      7122..8431
```

```
        Consensus quality: 147711 bases at least Q40
        Consensus quality: 155048 bases at least Q30
        Consensus quality: 158115 bases at least Q20
        Insert size: 170000: agarose-fp
        Insert size: 160133; sum-of-contigs
        Quality coverage: 3.3 in Q20 bases; agarose-fp
        Quality coverage: 3.5 in Q20 bases; sum-of-contigs
--------
      * NOTE: This is a 'working draft' sequence. It currently
      * consists of 24 contigs. The true order of the pieces
      * is not known and their order in this sequence record is
      * arbitrary. Gaps between the contigs are represented as
      * runs of N, but the exact sizes of the gaps are unknown.
      * This record will be updated with the finished sequence
      * as soon as it is available and the accession number will
      * be preserved.
      *        1     1374: contig of 1374 bp in length
      *     1375     1474: gap of     100 bp
      *     1475     2181: contig of 707 bp in length
      *     2182     2281: gap of     100 bp
      *     2282     2673: contig of 392 bp in length
      *     2674     2773: gap of     100 bp
      *     2774     5106: contig of 2333 bp in length
      *     5107     5206: gap of     100 bp
      *     5207     7656: contig of 2450 bp in length
      *     7657     7756: gap of     100 bp
      *     7757     9790: contig of 2034 bp in length
      *     9791     9890: gap of     100 bp
      *     9891    13844: contig of 3954 bp in length
      *    13845    13944: gap of     100 bp
      *    13945    16502: contig of 2558 bp in length
      *    16503    16602: gap of     100 bp
      *    16603    19777: contig of 3175 bp in length
      *    19778    19877: gap of     100 bp
      *    19878    23124: contig of 3247 bp in length
      *    23125    23224: gap of     100 bp
      *    23225    28282: contig of 5058 bp in length
      *    28283    28382: gap of     100 bp
      *    28383    31743: contig of 3361 bp in length
      *    31744    31843: gap of     100 bp
      *    31844    37479: contig of 5636 bp in length
      *    37480    37579: gap of     100 bp
      *    37580    44685: contig of 7106 bp in length
      *    44686    44785: gap of     100 bp
      *    44786    52369: contig of 7584 bp in length
      *    52370    52469: gap of     100 bp
      *    52470    60987: contig of 8518 bp in length
      *    60988    61087: gap of     100 bp
      *    61088    71201: contig of 10114 bp in length
      *    71202    71301: gap of     100 bp
      *    71302    81815: contig of 10514 bp in length
      *    81816    81915: gap of     100 bp
      *    81916    93418: contig of 11503 bp in length
      *    93419    93518: gap of     100 bp
      *    93519   107008: contig of 13490 bp in length
      *   107009   107108: gap of     100 bp
      *   107109   117125: contig of 10017 bp in length
      *   117126   117225: gap of     100 bp
      *   117226   129994: contig of 12769 bp in length
      *   129995   130094: gap of     100 bp
      *   130095   143708: contig of 13614 bp in length
      *   143709   143808: gap of     100 bp
      *   143809   162433: contig of 18625 bp in length.

FEATURES             Location/Qualifiers
     source          1..162433
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-611H22"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..1374
                     /note="assembly_fragment"
     misc_feature    1475..2181
                     /note="assembly_fragment
     misc_feature    2282..2673
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     misc_feature    2774..5106
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    5207..7656
                     /note="assembly_fragment"
     misc_feature    7757..9790
                     /note="assembly_fragment"
     misc_feature    9891..13844
                     /note="assembly_fragment"
     misc_feature    13945..16502
                     /note="assembly_fragment"
     misc_feature    16603..19777
                     /note="assembly_fragment"
     misc_feature    19878..23124
                     /note="assembly_fragment"
     misc_feature    23225..28282
                     /note="assembly_fragment"
     misc_feature    28383..31743
                     /note="assembly_fragment"
     misc_feature    31844..37479
                     /note="assembly_fragment"
     misc_feature    37580..44685
                     /note="assembly_fragment"
     misc_feature    44786..52369
                     /note="assembly_fragment"
     misc_feature    52470..60987
                     /note="assembly_fragment"
     misc_feature    61088..71201
                     /note="assembly_fragment"
     misc_feature    71302..81815
                     /note="assembly_fragment"
     misc_feature    81916..93418
                     /note="assembly_fragment"
     misc_feature    93519..107008
                     /note="assembly_fragment"
     misc_feature    107109..117125
                     /note="assembly_fragment"
     misc_feature    117226..129994
                     /note="assembly_fragment"
     misc_feature    130095..143708
                     /note="assembly_fragment"
     misc_feature    143809..162433
                     /note="assembly_fragment"
BASE COUNT    50320 a   29690 c   29259 g   50860 t   2304 others
ORIGIN

Query Match              0.5%;   Score 24;  DB 2;   Length 162433;
Best Local Similarity  100.0%;  Pred. No. 0.4;
Matches   24;  Conservative  0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   3803 gtaaaatatgtaataataa 3826
          ||||||||||||||||||||
Db  40394 GTAAAATATGTRATAATAATAA 40417

RESULT 11
AC018496/c
LOCUS       AC018496            163272 bp    DNA      linear   HTG 03-FEB-2000
DEFINITION  Homo sapiens chromosome 3p clone RP11-210B17, WORKING DRAFT
            SEQUENCE, 5 unordered pieces.
ACCESSION   AC018496
VERSION     AC018496.2  GI:6684201
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 163272)
  AUTHORS   Bao,J., Hu,S., Dong,W., Zhang,X., Wu,Q., Wang,H., Yang,X., Zhang,H.,
            Liu,B., Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y.,
            Niu,Y., Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H.,
            Liu,Y., Li,G., Li,C., Bao,Q., Wang,X., Wang,X., Song,L., Zhang,L.,
            Guo,D., Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L.,
            Feng,X., Yu,J. and Yang,H.
    TITLE   Chromosome 3p genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163272)
  AUTHORS   Li,T., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B., Bao,W.,
            Sun,Y., Wu,Q., Wang,H., Liu,N., Yang,X., Cheng,C., Wang,Y., Niu,Y., Qi,X.,
            Wang,X., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Li,G.,
            Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
            Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
            and Yang,H.
    TITLE   Direct Submission
  JOURNAL   Submitted (13-DEC-1999) Human Genomic Center, Institute of
            Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
            100101, P.R.China
  COMMENT   On Jan 9, 2000 this sequence version replaced gi:6553492.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     4841: contig of 4841 bp in length
            *                    gap of unknown length
            *     4842    26723: contig of 21882 bp in length
            *                    gap of unknown length
            *    26724    57119: contig of 30396 bp in length
            *                    gap of unknown length
            *    57120   104873: contig of 47754 bp in length
            *                    gap of unknown length
            *   104874   163272: contig of 58399 bp in length.
FEATURES             Location/Qualifiers
     source          1..163272
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3p"
                     /clone="RP11-210B17"
BASE COUNT    51500 a  30337 c  29864 g  51563 t      8 others
ORIGIN

Query Match
Best Local Similarity   100.0%;  Pred. No. 0.4;
Matches   24; Conservative   0; Mismatches   0;  Indels   0;  Gaps   0;

Qy   3803 gtaaaatatgtaataataataa 3826
          ||||||||||||||||||||||
Db  84958 GTAAAATATGTAATAATAATAA 84935

RESULT 12
AC026190
AC026190
LOCUS       AC026190      165077 bp    DNA     linear   HTG 29-MAY-2000
DEFINITION  Homo sapiens chromosome 3 clone RP11-372K9 map 3p, WORKING DRAFT
            SEQUENCE, 19 unordered pieces.
ACCESSION   AC026190
VERSION     AC026190.2  GI:8101204
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 165077)
  AUTHORS   Bao,W., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,J., Chen,J., Ding,H.,
            Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
            Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,
            Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,Y.,
            Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
            Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,R.,
            Wang,X., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
            Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
            Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
            Zhang,Z., Zhu,B., Yu,J. and Yang,H.
    TITLE   Chromosome 3p genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 165077)
  AUTHORS   Sun,W., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
            Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
            Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
            Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
            Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
            and Yang,H.
    TITLE   Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Human Genomic Center, Institute of
            Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
            100101, P.R.China
  COMMENT   On May 29, 2000 this sequence version replaced gi:7272007.
            Center: Beijing Center
            Center code: Beijing
            Website: http://hgc.igtp.ac.cn
            http://www.genomics.org.cn
            Contact: hgc@igtp.ac.cn
            --------    Project Information
            Center project name:i8 project
            Center clone name: RP11-372K9
            --------    Summary Statistics
            Sequencing vector: pUC18; 100% of reads
            Chemistry: Dye-terminator: ET 55% of reads
            Chemistry: Dye-terminator Big Dye; 45% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 151327 bases at least Q40
            Consensus quality: 159291 bases at least Q30
            Consensus quality: 164481 bases at least Q20
            Insert size: 158591: sum-of-contigs
            Quality coverage:    3.38x in Q20 bases:sum-of-contigs
            *
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 19 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1161: contig of 1161 bp in length
            *     1162     1261: gap of unknown length
            *     1262     2361: contig of 1100 bp in length
            *     2362     2461: gap of unknown length
            *     2462     4676: contig of 2215 bp in length
            *     4677     4776: gap of unknown length
            *     4777     7509: contig of 2733 bp in length
            *     7510     7609: gap of unknown length
            *     7610    10791: contig of 3182 bp in length
            *    10792    10891: gap of unknown length
            *    10892    15192: contig of 4301 bp in length
            *    15193    15292: gap of unknown length
            *    15293    20869: contig of 5577 bp in length
            *    20870    20969: gap of unknown length
            *    20970    27390: contig of 6421 bp in length
            *    27391    27490: gap of unknown length
            *    27491    35391: contig of 7901 bp in length
            *    35392    35491: gap of unknown length
            *    35492    42412: contig of 6921 bp in length
            *    42413    42512: gap of unknown length
            *    42513    50794: contig of 8282 bp in length
            *    50795    50894: gap of unknown length
            *    50895    57502: contig of 6608 bp in length
            *    57503    57602: gap of unknown length
```

```
RESULT 13
AC027763/c
LOCUS       AC027763      169770 bp      DNA      linear   PRI 01-FEB-2002
DEFINITION  Homo sapiens chromosome 17, clone RP11-530N7, complete sequence.
ACCESSION   AC027763
VERSION     AC027763.9  GI:18104884
KEYWORDS    HTG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 169770)
AUTHORS     Birren,B., Linton,L.L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens chromosome 17, clone RP11-530N7
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 169770)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,N., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,C., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,J., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 169770)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kanat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (01-FEB 2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jan 10, 2002 this sequence version replaced gi:15422062.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----------- Genome Center
```

```
        *          57603    67629: contig of 10027 bp in length
        *          67630    67729: gap of unknown length
        *          67730    79455: contig of 11726 bp in length
        *          79456    79555: gap of unknown length
        *          79556    91868: contig of 12313 bp in length
        *          91869    91968: gap of unknown length
        *          91969   102398: contig of 10430 bp in length
        *         102399   102498: gap of unknown length
        *         102499   117022: contig of 14524 bp in length
        *         117023   117122: gap of unknown length
        *         117123   135218: contig of 18096 bp in length
        *         135219   135318: gap of unknown length
        *         135319   165077: contig of 29759 bp in length.
FEATURES             Location/Qualifiers
     source          1..165077
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /map="3p"
                     /clone="RP11-372K9"
     misc_feature    1..1161
                     /note="assembly_name:Contig6"
     misc_feature    1262..2361
                     /note="assembly_name:Contig8"
     misc_feature    2462..4676
                     /note="assembly_name:Contig9"
     misc_feature    4777..7509
                     /note="assembly_name:Contig10"
     misc_feature    7610..10791
                     /note="assembly_name:Contig11"
     misc_feature    10892..15192
                     /note="assembly_name:Contig12"
     misc_feature    15293..20869
                     /note="assembly_name:Contig13"
     misc_feature    20970..27390
                     /note="assembly_name:Contig14"
     misc_feature    27491..35391
                     /note="assembly_name:Contig15"
     misc_feature    35492..42412
                     /note="assembly_name:Contig16"
     misc_feature    42513..50794
                     /note="assembly_name:Contig17"
     misc_feature    50895..57502
                     /note="assembly_name:Contig18
                     clone_end:SP6
                     vector_side:right"
     misc_feature    57603..67629
                     /note="assembly_name:Contig19"
     misc_feature    67730..79455
                     /note="assembly_name:Contig20
                     clone_end:T7
                     vector_side:right"
     misc_feature    79556..91868
                     /note="assembly_name:Contig21"
     misc_feature    91969..102398
                     /note="assembly_name:Contig22"
     misc_feature    102499..117022
                     /note="assembly_name:Contig23"
     misc_feature    117123..135218
                     /note="assembly_name:Contig24"
     misc_feature    135319..165077
                     /note="assembly_name:Contig25"
BASE COUNT    50675 a   30322 c   30406 g   51830 t    1844 others
ORIGIN

Query Match        0.5%;   Score 24;  DB 2;  Pred. No. 0.4;  Length 165077;
Best Local Similarity  100.0%;   Pred. No. 0.4;
Matches   24;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   3803 gtaaatatatgtaataataataa 3826
          ||||||||||||||||||||||
Db  50711 GTAAATATATGTAATAATAATAA 50734
```

```
          Center: Whitehead Institute/ MIT Center for Genome Research
          Center code: WIBR
          Web site: http://www-seq.wi.mit.edu
          Contact: sequence_submissions@genome.wi.mit.edu
          ---------- Project Information ----------
          Center project name: L8635
          Center clone name: 530_N_7
          ----------

FEATURES             Location/Qualifiers
source               1..169770
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-530N7"
                     /clone_lib="RPCI-11 Human Male BAC"
repeat_region        complement(1..321)
                     /rpt_family="MER66A"
repeat_region        complement(322..630)
                     /rpt_family="AluSx"
repeat_region        complement(631..670)
                     /rpt_family="MER66A"
repeat_region        749..773
                     /rpt_family="AT_rich"
repeat_region        complement(867..1176)
                     /rpt_family="AluJo"
repeat_region        complement(1437..1492)
                     /rpt_family="MIR"
repeat_region        1715..2085
                     /rpt_family="MLT1K"
repeat_region        complement(2140..2448)
                     /rpt_family="MSTB"
repeat_region        2456..2524
                     /rpt_family="AluSq"
repeat_region        complement(2709..3019)
                     /rpt_family="MLT1K"
repeat_region        3274..3324
                     /rpt_family="AluSp"
repeat_region        complement(3377..3784)
                     /rpt_family="(TTTA)n"
repeat_region        complement(3945..4235)
                     /rpt_family="MSTB"
repeat_region        complement(4565..4630)
                     /rpt_family="AluSx"
repeat_region        complement(5215..5528)
                     /rpt_family="L1MC3"
repeat_region        5546..5727
                     /rpt_family="L1ME1"
repeat_region        5736..6045
                     /rpt_family="AluSg"
repeat_region        6046..6092
                     /rpt_family="AluSx"
repeat_region        6107..6448
                     /rpt_family="L1PA16"
repeat_region        6449..6615
                     /rpt_family="AluSx"
repeat_region        complement(6616..6917)
                     /rpt_family="tigger3(Golem)"
repeat_region        6918..7108
                     /rpt_family="AluSg"
repeat_region        complement(7109..7404)
                     /rpt_family="tigger3(Golem)"
repeat_region        7405..7686
                     /rpt_family="AluY"
repeat_region        complement(7688..8054)
                     /rpt_family="tigger3(Golem)"
repeat_region        8055..8078
                     /rpt_family="L1MDa"
repeat_region        complement(8079..8316)
                     /rpt_family="(TTTTA)n"
repeat_region        complement(8318..8627)
                     /rpt_family="L1MDa"
repeat_region        complement(8647..9092)
                     /rpt_family="AluSx"

repeat_region        complement(9094..9916)
                     /rpt_family="L1MDa"
repeat_region        9917..10177
                     /rpt_family="L1ME1"
repeat_region        complement(10178..10672)
                     /rpt_family="AluSg"
repeat_region        complement(10673..10970)
                     /rpt_family="L1ME1"
repeat_region        complement(10971..11044)
                     /rpt_family="AluSx"
repeat_region        11116..11224
                     /rpt_family="L1ME1"
repeat_region        11393..11613
                     /rpt_family="CT-rich"
repeat_region        11721..12031
                     /rpt_family="MIR"
repeat_region        complement(12120..12263)
                     /rpt_family="AluSp"
repeat_region        12264..12292
                     /rpt_family="L1MA10"
repeat_region        complement(12293..12489)
                     /rpt_family="(ATG)n"
repeat_region        complement(12469..12490)
                     /rpt_family="L1MA10"
repeat_region        complement(12491..12802)
                     /rpt_family="L1M1"
repeat_region        complement(12803..13403)
                     /rpt_family="AluSx"
repeat_region        complement(13395..13598)
                     /rpt_family="L1M1"
repeat_region        complement(13599..13855)
                     /rpt_family="L1M1"
repeat_region        complement(13856..14134)
                     /rpt_family="MER70A"
repeat_region        complement(14135..14361)
                     /rpt_family="AluSg"
repeat_region        complement(14362..15829)
                     /rpt_family="MER70A"
repeat_region        15830..15850
                     /rpt_family="L1M1"
repeat_region        complement(15851..16503)
                     /rpt_family="(TTTG)n"
repeat_region        16504..16802
                     /rpt_family="L1M1"
repeat_region        complement(16803..17192)
                     /rpt_family="AluSg1"
repeat_region        complement(17194..17792)
                     /rpt_family="L1M1"
repeat_region        17793..17822
                     /rpt_family="L1M1"
repeat_region        complement(17823..18638)
                     /rpt_family="(TTTA)n"
repeat_region        complement(18612..18733)
                     /rpt_family="L1PA3"
repeat_region        complement(18839..19064)
                     /rpt_family="L1M3c"
repeat_region        complement(19161..19536)
                     /rpt_family="L1M3c"

Query Match
Best Local Similarity    0.5%;   Score 24;   DB 9;   Length 169770;
                       100.0%;   Pred. NO. 0.4;
Matches  24;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy  3819  aataataataaagcccacaag 3842
          ||||||||||||||||||||||
Db 16959  AATAATAATAAAGCCCACAAG 16936

RESULT 14
AC040977
LOCUS       AC040977        173878 bp    DNA    linear   HTG 05-SEP-2001
DEFINITION  Homo sapiens chromosome 17 clone RP11-589p10 map 17, *** SEQUENCING
            IN PROGRESS ***, 7 unordered pieces.
```

```
ACCESSION   AC040977
VERSION     AC040977.5  GI:15431174
KEYWORDS    HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 173878)
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens chromosome 17, clone RP11-589P10
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 173878)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 5, 2001 this sequence version replaced gi:14336616.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ------------------------  Genome Center
                  Center: Whitehead Institute/ MIT Center for Genome Research
                  Center code: WIBR
                  Web site: http://www-seq.wi.mit.edu
                  Contact: sequence_submissions@genome.wi.mit.edu
            ------------------------  Project Information
                  Center project name: L9393
                  Center clone name: 589_P_10
            ------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *          1   28017: contig of 28017 bp in length
            *      28018   28117: gap of    100 bp
            *      28118   31285: contig of 3168 bp in length
            *      31286   31385: gap of    100 bp
            *      31386   77861: contig of 46476 bp in length
            *      77862   77961: gap of    100 bp
            *      77962   91856: contig of 13895 bp in length
            *      91857   91956: gap of    100 bp
            *      91957  116594: contig of 24638 bp in length
            *     116595  116694: gap of    100 bp
            *     116695  132759: contig of 16065 bp in length
            *     132760  132859: gap of    100 bp
            *     132860  173878: contig of 41019 bp in length
FEATURES             Location/Qualifiers
     source          1..173878
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-589P10"
                     /clone_lib="RPCI-11 Human Male BAC"

BASE COUNT    49141 a  39705 c  39091 g  45256 t   685 others
ORIGIN

Query Match            0.5%;    Score 24;  DB 2;   Length 173878;
Best Local Similarity 100.0%;   Pred. No. 0.4;
Matches   24; Conservative  0; Mismatches  0; Indels  0; Gaps  0:

Qy   3819 aataataataataagccacaaag 3842
          ||||||||||||||||||||||||
Db  86976 AATAATAATAATAAGCCACAAAG 86999

RESULT 15
AC034192/c
LOCUS       AC034192              176217 bp    DNA     linear   PRI 11-FEB-2001
DEFINITION  Homo sapiens chromosome 3 clone RP11-416N8 map 3p, complete
            sequence.
ACCESSION   AC034192
VERSION     AC034192.4  GI:12745080
KEYWORDS    HTG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 176217)
AUTHORS     Xiong,H., Zhou,Y., Dong,H., Lin.W., Chen,B., Zhang,C., Zhang,Y.,
            Cai,Z., Ying,H.F., Wang,H., Gu,W., Zhu,G., Tu,Y., Zhang,X., Jia,J.,
            Shen,H., Zhang,D., Wu,C., Lu,G., Zhong,M., Jiang,H., Ren,S., Fu,G.,
            Chen,Z. and Huang,M.
TITLE       Chromosome 3p genomic sequence
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 176217)
AUTHORS     Li,L., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu.B., Bao,W.,
            Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y., Qi,X.,
            Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Li,G.,
            Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
            Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
            and Yang,H.
TITLE       Direct Submission
JOURNAL     Submitted (05-APR-2000) Human Genomic Center, Institute of
            Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
            100101, P.R.China
REFERENCE   3  (bases 1 to 176217)
AUTHORS     Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
            Dong,W., Fan,H., Feng,X., Gong,J., Gong,Q., Gu,X., Guo,D., Guo,Z.,
            He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F.,
            Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y.,
            Li,W., Li,Y., Luo,C., Luo,J., Niu,J., Qi,X., Song,H., Wang,J.,
            Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
            Wang,J., Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
            Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X.,
            Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
            Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
            Yu,J. and Yang,H.
TITLE       Direct Submission
JOURNAL     Submitted (11-FEB-2001) Human Genomic Center, Institute of
            Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
            100101, P.R.China
COMMENT     On Feb 11, 2001 this sequence version replaced gi:8101214.
            ------------------------Genome Center
                  Center:Beijing Center
                  Center code:Beijing
                  Website:http://hgc.igtp.ac.cn
                  http://www.genomics.org.cn
                  Contact:hgc@igtp.ac.cn
            ------------------------ Project Information
                  Center project name:1# project
```

```
          Center clone name: RP11-416N8
          ---------- Summary Statistics
          Sequencing vector: pUC18; 100% of reads
          Chemistry: Dye-terminator: ET 55% of reads
          Chemistry: Dye-terminator Big Dye; 45% of reads
          Assembly program: Phrap; version 0.990329
          Consensus quality: 179317 bases at least Q40
          Consensus quality: 179556 bases at least Q30
          Consensus quality: 179718 bases at least Q20
          Insert size: 176217; sum-of-contigs
          Quality coverage: 10.64x in Q20 bases;sum-of-contigs
          ----------
FEATURES          Location/Qualifiers
     source       1. .176217
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="3"
                  /map="3p"
                  /clone="RP11-416N8"
BASE COUNT    55712 a  32095 c  32227 g   56183 t
ORIGIN


Query Match          0.5%;   Score 24;   DB 9;   Length 176217;
Best Local Similarity   100.0%;   Pred.No. 0.4;
Matches   24;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   3803 gtaaaatatgtaataataataa 3826
          |||||||||||||||||||||||
Db 140840 GTAAAATATGTAATAATAATAA 140817


Search completed: September 19, 2002, 03:16:48
Job time: 17403 sec
```